

# VOICE STORING DEVICE AND VOICE ENCODING DEVICE

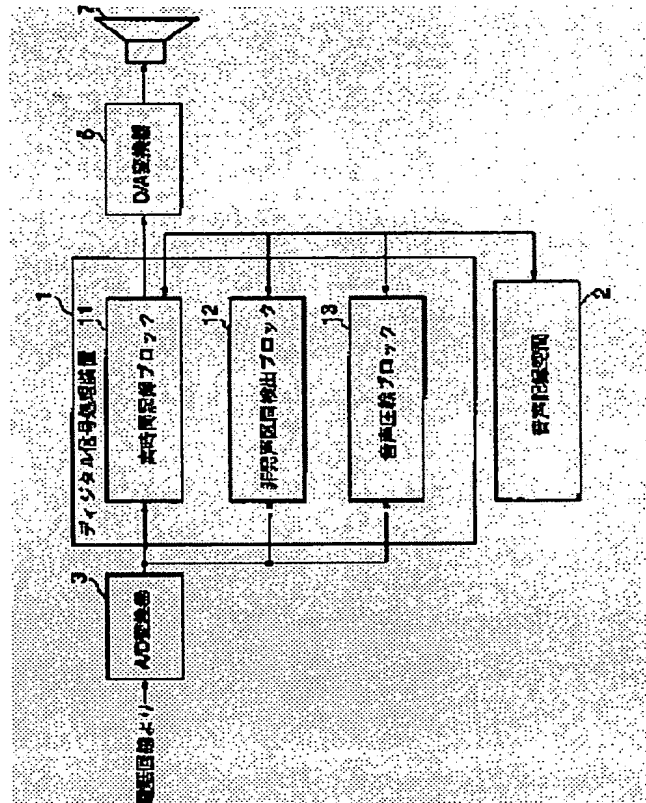
**Patent number:** JP2001053869  
**Publication date:** 2001-02-23  
**Inventor:** YOKOYAMA ATSUSHI  
**Applicant:** OKI ELECTRIC IND CO LTD  
**Classification:**  
 - international: H04M1/65; G10L19/00  
 - european:  
**Application number:** JP19990229005 19990813  
**Priority number(s):**

Report a data error here

## Abstract of JP2001053869

**PROBLEM TO BE SOLVED:** To clearly record voice in the case of the sufficient remaining quantity in a recording medium and to improve mission of voice even when the remaining quantity is reduced by compressing a voice signal among the stored ones, which requires encoding compression again, in accordance with compression and the importance of the voice signal.

**SOLUTION:** A frame importance degree is provided, which densely arranged to be  $\geq$  the number of usage possible encoding systems, a voice non-uttering section detecting block 12 is separated from a voice compression block 13 and the block 13 is operated as necessary. That is, when the sufficient remaining quantity exists in a voice recording space 2, the block 13 is not operated. When the quantity of a storage space is reduced, the block 13 starts operating. The voice compression block 13 scans a voice packet in the voice recording space 2 and voice is encoded and compressed in accordance with the frame importance degree which is set by a frame importance degree setting unit. Besides, an encoding ID is decided by the combination of the compression degree with the frame importance degree by referring to a reference table.



Data supplied from the esp@cenet database - Patent Abstracts of Japan



【特許請求の範囲】

【請求項1】 デジタル符号化した音声を蓄積する音声蓄積装置において、

音声信号を蓄積する蓄積手段と、

前記蓄積手段に蓄積された音声信号のうち再度符号化圧縮が必要なものについては圧縮の度合いと音声信号の重要度に応じて圧縮する音声圧縮手段とを有することを特徴とする音声蓄積装置。

【請求項2】 前記蓄積手段は、符号化方式ごとにパケット化した音声信号を蓄積することを特徴とする請求項1記載の音声蓄積装置。

【請求項3】 音声信号を実時間で前記蓄積手段に蓄積する実時間処理手段と、

前記実時間処理手段の処理とは独立したタイミングで非発音区間を検出する非発音区間検出手段とを有することを特徴とする請求項1記載の音声蓄積装置。

【請求項4】 前記非発音区間検出手段は、前記蓄積手段に蓄積された音声信号が発音区間である場合に、その前後の1又は複数の音声信号の重要度を増加させることを特徴とする請求項3記載の音声蓄積装置。

【請求項5】 前記音声圧縮手段は、前記蓄積手段の蓄積容量に応じて他の処理と並列又は時分割に音声信号を圧縮することを特徴とする請求項1記載の音声蓄積装置。

【請求項6】 前記蓄積手段より一度以上読み出されたか否かを記録する記録手段を有し、一度読み出された音声信号は前記重要度を低下させることを特徴とする請求項1記載の音声蓄積装置。

【請求項7】 一度読み出された前記音声信号が再度読み出された場合は、前記重要度を変更しないことを特徴とする請求項6記載の音声蓄積装置。

【請求項8】 音声信号の前記重要度を変更された場合、前記蓄積手段の蓄積容量に応じて前記重要度を変更するか否かを判断することを特徴とする請求項6記載の音声蓄積装置。

【請求項9】 前記音声圧縮手段は、前記蓄積手段に蓄積された音声信号を走査し、符号化圧縮のための量子化表を生成する量子化表生成手段を有することを特徴とする請求項1記載の音声蓄積装置。

【請求項10】 前記量子化表をパケット化して前記蓄積手段に蓄積し、前記蓄積手段から音声信号を読み出す場合に前記音声信号が前記量子化表を作成したのであれば、前記量子化表を用いて前記音声信号を復号することを特徴とする請求項9記載の音声蓄積装置。

【請求項11】 前記音声圧縮手段は、音声の復号に必須の符号と復号音声の音質を改善するための符号とに分けて符号化することを特徴とする請求項1記載の音声蓄積装置。

【請求項12】 LSPパラメータを量子化するLSP量子化器と、前記LSP量子化器では表現しきれない量

子化誤差成分を量子化するLSP多段量子化器と、線形予測残差からピッチ成分を求めるピッチ成分符号化器と、

励振源を多段のコードブック構成とする励振源符号化器と、

前記LSP量子化器、前記ピッチ成分符号化器、及び前記励振源符号化器の出力から復号に必須の符号を出力する必須符号出力器と、

前記LSP多段量子化器及び前記励振源符号化器の出力から復号音質を改善する符号を出力する音質改善符号出力器とを有することを特徴とする音声符号化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、デジタル符号化記録型の留守番電話装置や、その他のデジタル符号化音声記録装置に適用して好適な音声蓄積装置及び音声符号化装置に関するものである。

【0002】

【従来の技術】従来より、固定加入者電話端末機あるいは携帯電話端末機に搭載される留守番電話装置が利用されている。留守番電話装置は、端末機の利用者が電話に応答できない際に、内蔵する記録媒体（磁気テープや半導体メモリ）に電話発信者の用件を音声で記録するものであり、ボイスメールとも呼ばれる。

【0003】近年、デジタル信号処理を低価格で実現することの可能な半導体LSIが利用可能となったことにより、話者の音声 CELP（符号励振線形予測符号化）などの高能率符号化アルゴリズムで圧縮し、記録媒体に格納する方式の留守番電話装置が提案されている。これにより、同一の記録容量の媒体を用いた場合に、通常のPCM（パルス符号変調）方式やADPCM（適応差分PCM）方式などに比べてより多くの音声を記録できるようになり、また、半導体メモリでは録音された複数の用件の中から、特定の用件を選んで再生することを高速に実現可能である。

【0004】また、留守番電話装置が携帯電話端末機にも搭載されるようになってきている。しかし、同端末機に対する利用者の小型化の要求などから、端末機に搭載することのできる半導体メモリの容量には強い制限がある。このため、通常のCELP方式などを用いた方法では要求される音声記録時間を実現することができなかった。

【0005】このような理由から、留守番電話装置に音声区間検出器を組み合わせたものが実用化されている。この方法では、用件音声を符号化圧縮して媒体に記録する際に、話者の発声状態を監視している。これは、例えば音声ゲインをある閾値と比較することにより実現する。この比較結果を用いて、話者の発声と発声の合間の時間など比較的重要度の低い区間（非発声区間）においては音声の符号化および記録を中断し、非発声区間の継続時間の情報のみを媒体に記録する。結果として、符号

化効率を見かけ上増加させることが可能となり、記憶媒体の利用効率が向上する。

【0006】図2に從來から提案されている半導体メモリを記録媒体に用いたデジタル符号化記録型の留守番電話装置（以下、単に留守番電話と呼ぶ）の概略構成図を示す。A/D（アナログ／デジタル）変換器3は通話相手からの音声PCM信号に変換する。デジタル信号処理装置4はPCM信号を符号化圧縮および復号伸長する。半導体メモリ6は符号化された音声を記録する。D/A（デジタル／アナログ）変換器5はPCM信号の音声をアナログ信号に変換する。スピーカ7はアナログ音声を出力する。

【0007】

【発明が解決しようとする課題】しかし、音声区間の識別には音声ゲインや音声のピッチ（周期成分）情報などを用いて実現する方法が提案されているが、音声の背景雑音に対する信号対雑音（S/N）比が劣化した際に、その音声検出能力が低下しやすい。具体的には、非発声区間に隣接する発声区間において語頭および語尾音声を非発声区間と誤認識し、音声欠落しやすくなる。

【0008】本発明は、音声区間の検出を実時間で正確に行うことが困難であることに注目し、記録媒体の残容量に余裕がある場合は、用件音声を明瞭に記録することを可能とし、また、残容量が小さくなった場合においても音声の欠落、特に語頭音声の欠落を大幅に改善することを可能とするものである。

【0009】

【課題を解決するための手段】上記課題を解決するために、デジタル符号化した音声を蓄積する音声蓄積装置において、音声信号を蓄積する蓄積手段と、蓄積手段に蓄積された音声信号のうち再度符号化圧縮が必要なものについては圧縮の度合いと音声信号の重要度に応じて圧縮する音声圧縮手段とを設けるようにした。

【0010】また、上記蓄積手段は、符号化方式ごとにパケット化した音声信号を蓄積するようにした。

【0011】また、LSPパラメータを量子化する量子化するLSP量子化器と、LSP量子化器では表現しきれない量子化誤差成分を量子化するLSP多段量子化器と、線形予測残差からピッチ成分を求めるピッチ成分符号化器と、励振源を多段のコードブック構成とする励振源符号化器と、LSP量子化器、ピッチ成分符号化器、及び励振源符号化器の出力から復号に必須の符号を出力する必須符号出力器と、LSP多段量子化器及び励振源符号化器の出力から復号音質を改善する符号を出力する音質改善符号出力器とを設けるようにした。

【0012】

【発明の実施の形態】（1）第1の実施形態  
以下、本発明による蓄積型信号符号化装置に適用した第1の実施形態を図面を参照しながら詳述する。

【0013】（1-1）第1の実施形態の構成

図1は、デジタル信号処理装置1及び音声記録空間2の概要を示す。

【0014】音声記録空間2として、従来符号化圧縮した音声を記録するために利用していた半導体メモリをPCM符号を記録するために併用する。図3は、第1の実施形態を実現するための半導体メモリ内部の音声記録空間2の構成を示すものである。音声記録空間2には、複数のパケット201が蓄積される。各パケットは可変長のヘッダ210と可変長のペイロード220から構成され、デジタル信号処理装置1内又は別個に設けられるパケット生成手段（図示せず）で生成される。

【0015】音声は20ミリ秒単位の大さき（フレームと呼ぶ）に分割され符号化される。各々のフレームは各パケットに格納される。ここで、各パケット（後述する空パケットおよび非発声パケットなど、音声を格納していないものを除く）が対応する音声の時間は常に20ミリ秒であるが、その物理的な符号の大さき（バイト数）は可変である。各パケットはPCM符号で表された音声信号の符号であったり、またCELP方式で圧縮された符号であったりする。

【0016】留守番電話装置は呼を着信する毎に、端末の利用者が呼に応答しなければ要件を音声記録空間2に記録する。音声記録空間2には、その容量が許す限り複数の要件を記録できる。以下、個々の要件を音声要件という。

【0017】パケット201のヘッダ210は図4に示すように、複数の情報フィールドから構成される。この情報フィールドのうち1つは4ビットのIDフィールド211であり、パケットの型（タイプ）を示すものである。本実施形態では、図5に示すパケットの型を有する。

【0018】空パケット（ID=0）は、音声記録空間2の未使用領域に対応する。後述するように音声信号の圧縮および再圧縮などにより生じるメモリの空き領域も、空パケットにより実現される。空パケットのヘッダには、パケットの大さき（バイト長）を示す情報フィールドがあり、ペイロードの中身は、例えば、不定値データの集まりである。

【0019】PCM符号パケット（ID=1）は、PCM符号化で表される音声信号を格納している。8KHz標準化の16ビット量子化PCMの場合、パケットのペイロードの大さきは320バイトとなる。

【0020】8k bps（キロビット毎秒）CELP符号パケット（ID=2）は、符号化速度8キロビット毎秒のCELP方式で符号化された音声信号を格納している。パケットのペイロードの大さきは20バイトとなる。

【0021】4k bps（キロビット毎秒）CELP符号パケット（ID=3）は、符号化速度4キロビット毎秒のCELP方式で符号化された音声信号を格納してい

る。パケットのペイロードの大きさは10バイトとなる。

【0022】非発声パケット(ID=4)は、非発声区間を表している。ペイロードには、非発声区間の時間(フレーム数)が格納されている。パケットのペイロードの大きさは1バイトである。257フレーム以上の非発声区間を表すには、複数の非発声パケットに分割する。

【0023】終端パケット(ID=5)は、個々の音声要件の末尾を表すための指標である。ペイロードは0バイトである。

【0024】各パケット(音声を格納していないものを除く。以下、音声パケットと呼ぶ)のヘッダに、さらにフレームの重要度を示す4ビットの重要度フィールド212を設ける。フレーム重要度は0から15で表される数値であり、数値が大きいくほど当該パケットに含まれる音声情報がより重要であることを意味する。発声区間の音声は非発声区間の音声に比べて重要であり、また、一度も再生されておらず、端末機の利用者が内容を確認していない用件音声は、一度以上再生された用件音声に比べて重要である。なお、これら重要度の基準は、留守番電話装置が使用される用途など、様々な因子によって別途決定することができる。

【0025】なお、フレーム重要度0は例外的な意味を持ち、当該パケットの音声フレームの重要度が、未判定であることを示す。

【0026】ディジタル信号処理装置1は、実時間記録ブロック11、非発声区間検出ブロック12、及び音声圧縮ブロック13から構成される。各ブロックは並列に動作し、動作上の優先順位は、実時間記録ブロック11が最も高く、次に非発声区間検出ブロック12、最後に音声圧縮ブロック13の順となっている。なお、処理ブロックはハードウェア回路又はソフトウェアプログラムのもちらでも実現可能である。

【0027】実時間記録ブロック11は、用件音声を音声記録空間2に保存する。本実施形態では、音声をPCM符号パケットとして音声記憶空間2に格納していく。このとき、フレーム重要度は0(未判定)としておく。

【0028】非発声区間検出ブロック12は、ある音声フレームが発声区間であるか非発声区間であるかに応じてフレーム重要度を判定し、記録していく。本実施形態では、重要度は1以上かつ15以下の値の選択肢の中から選ばれる。選択方法としては、例えば、入力パワーを所定の閾値と比較することにより選択する。そして、最も発声区間らしい音声フレームには15の値が設定され、最も非発声区間らしい音声フレームには1の値が設定される。

【0029】次に、図6を用いて非発声区間検出ブロック12を説明する。ピッチ検出器121は、音声パケット内のフレームの音声信号を用いて、音声に含まれるピ

ッチ成分の周波数(周期)と、ピッチ予測ゲインを求める。パワ演算器122は、音声フレームのパワを求める。フレームS/N比演算器123は、音声フレームパワと背景雑音レベルの推定値を比較しフレーム毎のS/N比を求める。背景雑音レベル推定器124は、求めた音声パワを用いて背景雑音の最低ゲインレベルを追跡する。音声判定器125は、ピッチ周波数およびピッチ予測ゲイン、フレームS/N比を用いて当該フレームの音声らしさを判定する。フレーム重要度設定器126は、音声判定器125の出力を用いて、音声フレームのパケットのフレーム重要度情報を設定する。低速変化S/N比推定器127は、フレームS/N比演算器123の出力をローパスフィルタ処理することで、その低速変化成分を出力する。先読み分析フレーム数演算器128は、低速変化S/N比から後述する先読みフレーム数を求める。前方後方ハングオーバー回路129は、前述の先読み分析フレーム数分だけ遅れて動作し、その過去の音声フレームのフレーム重要度が一定の閾値以上である場合に、その時間前後のフレームのフレーム重要度を増加させる。

【0030】次に、図8を用いて音声圧縮ブロック13を説明する。音声記録空間残量監視器1301は、音声記録空間2の残量を常時監視する。音声符号化器1302はPCM符号の音声信号をCELP方式で圧縮する。圧縮参照表1303は、所望の圧縮度合いとフレームの重要度に応じて符号化IDを決定する。圧縮度合レジスタ1304は、現在の所望の圧縮度合いを保持する。音声記録空間2を収める半導体メモリと同様に、このレジスタは電池などで記憶が保持される不揮発半導体メモリなどで構成される。

【0031】図9はフレーム重要度設定のための構成の一例を示す。再生指示回路1210は、留守番電話装置に保持されている用件音声の再生を指示する。用件音声指標テーブル1211は、各用件音声毎に当該用件音声一度以上再生されたか否かを記録する。音声要件が一度以上再生されていれば、当該音声要件の全体のフレーム重要度を引き上げることとする。

【0032】(1-2)第1の実施形態の動作  
以下、第1の実施形態の蓄積型信号符号化装置の動作を説明する。

【0033】実時間記録ブロック11は、音声をPCM符号パケットとして音声記憶空間2に格納していく。このとき、フレーム重要度は0(未判定)としておく。

【0034】非発声区間検出ブロック12は、実時間記録ブロック11が処理しているフレームよりも、ある可変時間だけ遅れて動作する。これは、発声区間と非発声区間を高精度で識別するためには、実時間処理が困難であるためであり、遅延時間については、音声の背景雑音に対するS/N比に応じて決定される。例えば、音声の背景雑音に対するS/N比が十分高い場合は、20フレ

ームだけ遅れて動作する。つまり、実時間記録ブロック11がn番目のフレームをバケット化しているとき、非発声区間検出ブロック12はn-20番目のフレームを処理する。また、音声の背景雑音に対するS/N比が低い場合は、判定が難しくなることから、さらに多くのフレーム分遅れて動作する。例えば、S/N比が5dBのときは50フレーム遅れて動作する。一般に、S/N比が低いほど、発声区間と非発声区間の識別は困難である。特に、非発声区間直後の発声区間（すなわち語頭）では、多くの音声フレームを先読みして分析することにより、その語頭を正しく発声区間と識別できる。このため、非発声区間検出ブロック12は、S/N比が低いほど大きく遅れて非発声区間を検出するようにしている。

【0035】次に、図6の非発声区間検出ブロック12の構成例を基に動作を説明する。ピッチ検出器121では、音声記録空間2中の音声バケットから取り出した入力音声の離散値信号を線形予測し、その残差信号を処理することで、ピッチ周波数およびピッチ予測ゲインを求め、音声判定器125に出力する。これらのパラメータは、入力音声の人間らしさの良し悪しを表すものである。パワ演算器122では、音声記録空間2中の音声バケットから取り出した入力音声の離散値信号の標本毎の自乗和を積算することで、音声フレーム単位のパワ値を求め、フレームS/N比演算器123及び背景雑音レベル推定器124に出力する。フレームS/N比演算器123では、パワ演算器122の出力と背景雑音レベル推定器124の出力の差を求めることで、音声フレーム毎のS/N比を求め、音声判定器125及び低速変化S/N比推定器127に出力する。背景雑音レベル推定器124では、パワ演算器122の出力から入力音声の最低パワレベル値を時間軸上で追跡することで、音声に重畳している背景雑音のレベルを推定し、その結果をフレームS/N比演算器123に出力する。音声判定器125では、ピッチ検出器121から出力されるピッチ周波数及びピッチ予測ゲイン、フレームS/N比演算器123から出力されるフレームS/N比の情報を用いて、当該フレームの信号の音声らしさを決定し、フレーム重要度設定器126に出力する。この値は発声区間では高くなり、非発声区間では小さな値となる。一般に、フレームS/N比が高いとき、それは発声区間と考えられる。しかしフレームS/N比が高くとも、ピッチ周波数が一定区間（人間の発声周波数の範囲内）から外れていたり、ピッチ予測ゲインが低いときは、人間の発声らしくない音響信号であると考えられる。これらの情報を総合評価し、音声らしさの情報を出力する。フレーム重要度設定器126では、音声判定器125の出力を1から15の値に正規化することで、フレーム重要度パラメータを決定する。この値は音声記録空間2の音声バケットのヘッダ中にあるフレーム重要度情報フィールドに格納される。低速変化S/N比推定器127では、フレームS

/N比演算器123の出力をローパスフィルタによって平滑化することで、低速度に変化するS/N比推定情報を先読み分析フレーム数演算器128に出力する。先読み分析フレーム数演算器128では、低速変化S/N比推定器127から出力される低速変化S/N比推定情報に基づいて、前方後方ハングオーバー回路129が先読み分析するフレーム数を決定する。推定S/N比が高い場合（例えば40dB以上の場合）、分析フレーム数を20フレームとする。逆にS/N比が低い場合（5dB程度）分析フレームを50フレームと長くする。このように、推定S/N比に逆比例して先読み分析フレーム数を決定する。前方後方ハングオーバー回路129では、上述のフレーム重要度設定器126がフレーム重要度情報を設定したフレームに対して、あるフレーム数だけ遅れて動作する。このフレーム数は、先読み分析フレーム数演算器128で求めた数である。前方後方ハングオーバー回路129は、注目している音声フレームのフレーム重要度が一定以上（ここでは12以上とする）であるとき、前後双方向に対して先読みフレーム数分だけ、フレーム重要度を調整する。具体的にはフレーム重要度に4を加算する。ただし、11を超えない値とする。（11とは、上記値12から1を減じた値である。）これにより、音声の背景雑音に対するS/N比が小さいとき、発声区間の前後（語頭および語尾）のフレーム重要度が小さくなることで、発声区間が非発声区間として誤判定されることを防止することができる。言い換えれば、同語頭あるいは語尾部分の欠落を防止する働きを持たせることができる。図7に、本動作をグラフで示す。グラフは、横軸が時間（バケット）であり、縦軸が当該バケットの音声のフレーム重要度である。ハッチ部分は、前方後方ハングオーバー回路によってフレーム重要度が増加されたことを示している。これにより、語頭において適切にフレーム重要度が修正されている様子が分かる。

【0036】半導体メモリ中の音声記録空間2の残容量が十分に大きい場合（例えば、これを残容量の閾値と比べることで判定する）は、これ以上の処理は行わない。つまり、記録空間に十分に余裕がある場合は、用件音声は全てPCM符号で記録されることになり、用件音声の再生時には、十分な音質が保持されている。

【0037】記憶空間の残容量が小さくなった場合は、音声圧縮ブロック13が動作を開始する。同ブロックは音声記録空間2中の音声バケットを走査し、フレーム重要度設定器126が設定したフレーム重要度に応じて音声の符号化圧縮する。同ブロックは図11に一例を示す参照表をひくことで、圧縮度合いとフレーム重要度の組み合わせから、符号化IDを決定する。例えば、最初（全バケットがPCM符号バケットである場合）は、フレーム重要度に関わらず全ての音声フレームを8kbps CELP（ID=2）で圧縮していく。それでも記憶空間の残容量が切迫してきた場合、フレーム重要度が

小さいものから4 kbps CELP (ID=3) で再圧縮する。さらに記憶容量が切迫したときには、フレーム重要度が4以上のものは全て4 kbps CELP (ID=3) で符号化圧縮され、それ以外は非発声パケット (ID=4) に変換される。最終的に、記憶空間が完全に一杯になったときは、用件録音の一部が削除されるまで新たな用件音声の録音ができなくなる。

【0038】以下、図8の音声圧縮ブロック13の動作を図12のフローチャートを用いて説明する。音声記録空間2の残量に余裕がある場合、音声圧縮ブロック13は動作しない。しかし音声記録空間2残量監視器1301が、音声記録空間2の残量が一定の閾値 (例えば20%) よりも小さくなったことを検出したとき、音声圧縮ブロック13は動作を開始する。音声符号化器1302は、圧縮度合レジスタ1304の値を読み出す (ステップS1)。音声符号化器1302は、音声記録空間2中の音声フレームを収めたパケットを先頭より読み出していく (ステップS2)。そして、その音声フレームのフレーム重要度と上述の圧縮度合レジスタ1304を用いて圧縮参照表1303を検索し、当該フレームをいずれの符号化方式のパケット型に変換すべきかを決定する (ステップS3)。圧縮参照表1303の内容については後述する。音声符号化器1302は、圧縮参照表1303で求めたパケット型が、現在のパケット型と異なるとき、当該パケットの音声信号を新たなパケット型の符号化方式で圧縮する。例えば、元のパケット型がPCM符号であり、新たなパケット型が8 kbps CELPである場合、パケットのペイロードから取り出したPCM符号を8 kbpsの符号化レートでCELPアルゴリズムで符号化圧縮する。そして、圧縮後の符号を再度パケットとして構築し、当該パケットのヘッダについてを再設定すると共に音声記録空間2に保管する (ステップS4)。ここで、符号化圧縮により新パケットは元パケットに比べて小さな記憶空間しか必要としないため、パケットの後半部分が空き空間となる。このままでは小さく分割された空パケットが散在してしまうため、図10に示す音声記録空間2を読み書きするための2つの指標 (読み出し指標230および書き込み指標231) を設ける。圧縮前の元のパケットは読み出し指標231を用いて読み出す。圧縮後の新パケットは、書き込み指標230を用いて書き込む。新パケットが元パケットよりも大きくなることはないため、書き込み指標230が読み出し指標231よりも後 (図における「右」) にくることではない。各指標としては、例えば、ポインタあるいはカウンタ (計数器) を使用する。なお、空き空間を空パケットとして設定するには、空パケットとしてのID=0をIDフィールドに、さらに同パケットの大きさをパケット長フィールドにそれぞれ書き込む。以上の処理により、原則として空パケットは音声記録空間2の後半にまとめられる。例外としては、音声圧縮ブロック13が

動作中に新たな用件音声の記録をする必要が生じた場合などに、止むを得ず音声圧縮ブロック13の動作を中断する場合がある。この場合、音声記憶空間2の途中に空パケットが存在することになる。このような空パケットは、例えば、デジタル信号処理装置1に空パケット収集ブロックを追加し、そのブロックの働きによって収集する。この動作例としては、計算機工学で用いられる一般的な技法である「ごみ集め」 (garbage collection) があげられる。なお、音声記憶空間2を順読み出しする場合は、空パケットを検出したとき、そのヘッダを参照することで空パケットの大きさを調べ、その大きさだけ音声記憶空間2を読み飛ばすことも可能である。ただし、小さな空パケットが分散していると、後で再び音声要件を記録する際に都合が悪いため、可能な限り空パケット収集ブロックを動作させておくことが望ましい。以上の処理を、音声記録空間2中の全音声パケットに対して繰り返す (ステップS5、S6)。しかし、最後まで処理を終了しても、音声記録空間2の残量が残量閾値よりも大きくならなかった場合、以下の処理を実行する。まず、圧縮度合レジスタ1304から読み出した値を1だけ増加させる (ステップS9)。そして、上述の音声パケット圧縮処理を、音声記録空間2の先頭からやり直す。最後に、音声記録空間2の残量が残量閾値よりも大きくなったとき、そのときの圧縮度合の値をレジスタ1304に格納し、音声圧縮ブロック13を終了する (ステップS7、S8)。

【0039】圧縮参照表1303の例を、図11に示す。なお、用件音声の記録動作中 (すなわち、実時間記録ブロック11の動作中) に音声記録空間2の残量が閾値よりも小さくなったとき、実時間記録ブロック11の動作に並行して音声圧縮ブロック13を動作させることができる。このとき、実時間記録ブロック11を最優先で動作させるが、音声圧縮ブロック13は、デジタル信号処理装置1の能力の範囲内で音声圧縮し、音声記録空間2の残量を拡大する。この結果、音声圧縮ブロック13のパケット指標が実時間記録ブロック11のパケット指標に追いついた場合は、音声圧縮ブロック13は実時間記録ブロック11に信号を伝達し、実時間記録ブロック11のパケット指標を空パケットの先頭位置まで復元させる。これにより、実時間記録ブロック11は音声記録空間2の残り空間を拡大することができる。

【0040】次に、図9のフレーム重要度の設定動作を説明する。留守番電話装置の利用者が、再生指示回路1210を介して、録音された用件音声の中から任意の用件の再生を指示すると、デジタル信号処理装置1は半導体メモリ中の音声記録空間2の中から、指示された用件音声の先頭パケットから順にパケットを取り出し、そのヘッダの符号化IDモードに従って音声復号し、再生する。その際、用件音声を完全に再生し終わると、デジタル信号処理装置1は、用件音声に属する全パケッ

トのヘッダを走査し、フレーム重要度情報を一定値（ここでは4）だけ減少させる。その後、同一の用件音声を複数回再生しても繰り返しフレーム重要度を減少させることのないよう、用件音声指標テーブル1211の中の用件指標の再生済みを示すフィールドに印をつけておく。一度フィールドに印が付けられた用件音声を繰り返し再生しても、フレーム重要度情報を再度減少させることをしない。ここではフレーム重要度の変更しか行わないが、後に音声記録空間2の残量が減少することで音声圧縮ブロック13が動作を開始すると、そのとき初めて、一度以上再生された用件音声について再圧縮が実施され、音声記録空間2の残量が拡大される。

#### 【0041】（1-3）第1の実施形態の効果

以上説明した第1の実施形態によれば、利用可能な符号化方式の数以上に細分化したフレーム重要度を設けると共に、非発声区間検出ブロック12と音声圧縮ブロック13を分離させ、必要とときだけ音声圧縮ブロック13を動作させることで、同一のフレーム重要度を持つパケットであっても、半導体メモリ中の音声記録空間2に余裕がある間は、もっとも最近に記録された用件音声についてはPCM符号のまま音声を記録しておくことが可能となり、記録された用件音声の音質を可能な限り高く保持することができる。

【0042】また、録音された用件音声が増加してきた場合、その用件音声の総計録音時間に応じて最低限必要なだけ、音声を高能率の符号化方式（CELPなど）で圧縮することができる。音声圧縮処理は非通話時に行うため、通話中のデジタル信号処理装置1に大きな負荷をかけることなく、音声を圧縮することができる。圧縮参照表を用いることにより、再圧縮が必要となるパケット数を少なく押さえることが出来、よりデジタル信号処理装置1の負荷を軽減することも可能である。非発声区間の判定に際しては、十分な時間の音声を読み出して分析するため、判定の精度が向上し、非発声区間直後の、音声語頭の記録欠落を防止することができる。

【0043】さらに、一度再生された用件音声に属するフレーム重要度を減少させることで、既に内容を確認済みの用件音声が消費する音声記録空間2を削減することが可能になる。一方、必要以上にフレーム重要度を減少させると、フレームが非発声パケットとして扱われるため、最終的に音声として復号不可能となることから、用件音声指標テーブル中の再生済みを示すフィールドに印をつけることで、同一の用件音声を繰り返し再生しても、必要以上にフレーム重要度を減少させることを防止できる。また、フレーム重要度情報の変更と実際の音声圧縮ブロック13を分離させることで、フレーム重要度に差のあるパケットが併存しても、半導体メモリ中の音声記録空間2に余裕がある間は、必要以上に音声の符号化圧縮を行わないことにより、記録された用件音声の音質を保持することが可能になる。

#### 【0044】（2）第2の実施形態

次に、本発明による蓄積型信号符号化装置に適用した第2の実施形態を図面を参照しながら詳述する。第2の実施形態は、第1の実施形態の音声圧縮ブロック13を改良したものであり、音声圧縮ブロック13を中心に説明する。

#### 【0045】（2-1）第2の実施形態の構成

図13は、第2の実施形態を実現するための音声圧縮ブロック13の構成を示すものである。音声復号器1305は、既に符号化圧縮されている音声パケットを復号伸長する。線形予測分析器1307は、音声のフォルマント特性を分析する。スイッチ1306は音声パケットの内容に応じて、線形予測分析器1307に与える信号を切り替える。LSP変換器1308は、線形予測パラメータ（LPCパラメータ）を線スペクトル対パラメータ（LSPパラメータ）に変換する。LSP量子化表生成器1309は、LSPパラメータ群の時間軸上の変化を走査分析し、LSPを離散符号に量子化するための量子化表を生成する。ピッチ分析器1310は、線形予測分析後の残差信号から、ピッチ成分の周波数およびピッチ予測ゲインを分析する。ピッチ分析パラメータ量子化表生成器1311は、ピッチ分析器で求めたパラメータ群の時間軸上の変化を走査分析し、ピッチ分析パラメータ群を離散符号に量子化するための量子化表を生成する。

【0046】ここで、線形予測分析器1307、LSP変換器1308、及びピッチ分析器1310は、実施形態1で説明した音声圧縮ブロック13における音声符号化アルゴリズムでも用いられるものであり、兼用することが可能である。

【0047】また、記録空間2中のパケットのヘッダに収められるIDフィールドのID番号として、新規にID=6を量子化表パケットとして追加する。

【0048】（2-2）第2の実施形態の動作第2の実施形態の音声圧縮ブロック13は、用件音声の一件単位に、音声の符号化圧縮アルゴリズムで使用する量子化表の群を、個別に生成する。CELPに代表される音声符号化方式では、音声の分析結果から得られる情報を量子化表を用いて離散符号に置き換えるが、この量子化表は汎用に設計されており、通常、あらゆる話者の音声の性質や言語の性質に対応できるようになっている。しかし、留守番電話の一用件音声に限ってみると、その短時間の特定話者かつ特定言語、特定性別の音声の性質においては、そのLSPパラメータやピッチ分析パラメータの分散量は制限されている。このため、一用件音声だけに特化された量子化表群を生成し使用することにより、汎用の量子化表を用いる場合に比べて、少ない符号情報量でも同等の音質で符号化できる。なお、量子化表の生成には多くの演算処理量を必要とするが、本特許の留守番電話装置の音声圧縮ブロック13は、用件音声の録音



時に実時間で動作する必要がないため、問題にはならない。

【0049】音声圧縮ブロック13では、デジタル信号処理装置1は音声記録空間2上の音声を符号化圧縮する際に、一用件音声単位で量子化表群を生成する。デジタル信号処理装置1は音声記録空間2から音声パケットを読み込むが、その内容が既に8k bps CELP方式で符号化されている場合、音声復号器1305により音声パケットを復号し、PCM符号の音声信号に変換する。スイッチ1306は、音声記録空間2から読み出した音声パケットがPCM符号パケットである場合に、そのPCM符号をそのまま線形予測分析器1307に送出する。線形予測分析器1307とLSP変換器1308は、音声記録空間2中に収められた用件音声に属するパケットを順次読み出し、LSPパラメータを出力する。LSP量子化表生成器1309は、LSPパラメータを順次取り込みながら、細かな区間を有するヒストグラム上に、その値の出現頻度を記録していく。同様に、ピッチ分析器1310とピッチ分析パラメータ量子化表生成器1311は、ピッチパラメータ群の出現頻度を記録していく。デジタル信号処理装置1は一用件音声を走査分析し終わると、LSPパラメータ量子化表およびピッチ分析パラメータ量子化表を生成する。その後、デジタル信号処理装置1は、生成した量子化表を用いて音声の符号化圧縮を開始する。その際、音声パケットを音声記録空間2に記録する前に、生成した量子化表をパケットに格納して、音声記録空間2に記録する。この際、パケットのヘッダのIDフィールドには、量子化表が収められていることを示すID=6が設定される。

【0050】次に、用件音声を再生する場合を説明する。デジタル信号処理装置1は音声記録空間2からパケットを取り出した際、それが量子化表のパケットならば、圧縮された音声符号を復号する際に用いる量子化表を取り出した量子化表で置き換える。その後、順次パケットを取り出しながら音声を復号、再生していく。

【0051】(2-3) 第2の実施形態の効果  
以上説明した第2の実施形態によれば、音声の符号化圧縮アルゴリズムで用いる量子化表群を、特定の用件音声専用で作成することができ、用件音声を発声した特定話者、特定言語、特定性別に最適化した量子化表群が得られ、汎用の量子化表群を用いる場合に比べて、同一の情報符号化速度であるならば、より音質の高い符号化が可能になる。

【0052】また、量子化表の生成には多くの演算処理が必要であるが、本実施形態の音声圧縮ブロック13では、量子化表を用件音声の記録時に実時間で生成する必要がないため、低い能力の信号処理プロセッサを用いても実現することが可能である。

【0053】(3) 第3の実施形態  
次に、本発明による蓄積型信号符号化装置に適用した第

3の実施形態を図面を参照しながら詳述する。第3の実施形態は、第1の実施形態における8k bps CELP符号化と4k bps CELP符号化を、階層化符号化の構成としたものである。

【0054】(3-1) 第3の実施形態の構成

以下、第3の実施形態の蓄積型信号符号化装置の構成を説明する。

【0055】図14は、第3の実施形態を実現するための階層化符号化の構成を示すものである。線形予測分析器1001及びLSP変換器1002は、既に説明した線形予測分析器1307及びLSP変換器1308と同等のものである。LSP量子化器1003は、LSP変換器によって得られるLSPパラメータを量子化する。LSP多段量子化器1004は、LSP量子化器1003の量子化誤差をさらに精密に量子化する。ピッチ分析符号化器1005は、線形予測残差におけるピッチ成分を求め、ピッチ合成フィルタを構成する。励振源符号化器1006は、ピッチ成分を取り除いた残差信号を用いて励振信号を生成符号化する。ここで、励振源符号化器1006は励振源を多段のコードブック構成で表すことで、復号音質に大きく影響する符号と、復号音声を若干改善するための符号の2種類を出力する。必須符号出力器1007は、LSP量子化器1003の出力、ピッチ分析符号化器1005の出力、又は励振源符号化器1006の出力から復号音質に大きく影響する符号をバックする。必須符号出力器1007の出力は4k bpsである。音質改善符号出力器1008は、LSP多段量子化器1004の出力又は励振源符号化器1006の出力から復号音質を改善する符号をバックする。音質改善符号出力器1008の出力は4k bpsである。

【0056】(3-2) 第3の実施形態の動作

線形予測分析器1001は、音声パケットから取り出した音声(離散値サンプル群)を線形予測分析することで線形予測パラメータを出力する。LSP変換器1002は、線形予測分析器1001から出力された線形予測パラメータを線スペクトル対(LSP)パラメータに変換する。LSP量子化器1003とLSP多段量子化器1004は、共にLSP変換器1002から出力されたLSPパラメータを離散符号に量子化するものであるが、LSP量子化器1003はLSPパラメータを粗く量子化し、LSP多段量子化器1004はLSP量子化器1003では表現し切れない微細な誤差成分を精密に量子化する。ピッチ分析量子化器1005は、音声パケットから取り出した音声及びLSP量子化器1003から出力された量子化値より線形予測残差を算出してピッチ成分を求め、ピッチ成分を取り除いた残差成分を出力する。励振源符号化器1006は、ピッチ分析量子化器1005から出力されたピッチ成分が取り除かれた残差成分をコードブック探索により符号化する。このとき、複数の励振パルスに、音質に寄与する度合いに応じて優

先順位を設ける。優先度の高い励振パルスに対応するコードブックの指標と、優先度の低い励振パルスに対応するコードブックの指標を別々に出力する。必須符号出力器1007は、LSP量子化器1003の出力、ピッチ分析符号化器1005の出力、又は励振源符号化器1006の出力から復号音質に大きく影響する符号をバックして、4kbpsの符号として出力する。音質改善符号出力器1008は、LSP多段量子化器1004の出力又は励振源符号化器1006の出力から復号音質を改善する符号をバックして、4kbpsの符号として出力する。

【0057】以上により、音声を復号する上で不可欠の符号(必須符号)と、復号には不可欠ではないものの復号音質の音質を向上させることのできる符号(音質改善符号)の2種類を出力する。前者の符号は符号化速度4kbpsであり、前者と後者の符号を合わせた符号化速度は8kbpsとなる。

【0058】さらに、本実施形態により、実施形態1の音声圧縮ブロック13を改良することができる。音声圧縮ブロック13は、音声フレームを8kbps CELPで圧縮する必要がある際に、本実施形態の必須符号と音質改善符号を併せてパケットに格納する。パケットのヘッダIDは8kbps CELP (ID=2)である。後に音声圧縮ブロック13が、8kbps CELPで圧縮された音声フレームを4kbps CELPで再圧縮する必要性が生じた際は、8kbps CELP方式による音声の復号と4kbps CELP方式の音声の符号化を行う代わりに、パケットの中から音質改善符号の4kbps符号のみを削除し、パケットのヘッダIDをID=2からID=3に変更するだけで良い。

【0059】(3-3)第3の実施形態の効果  
以上説明した第3の実施形態によれば、8kbps CELPの符号が納められたパケットを4kbps CELPに再圧縮する必要がある場合、復号および再符号化の代わりにパケットの一部を削除するだけで済むため、信号処理プロセッサの処理量を大幅に削減することが可能になる。

【0060】また、復号および再符号化の処理は、一度量子化された分析結果を用いて復元された音声を再分析し、さらに再量子化をしなくてはならないため、最初から4kbps CELP方式で符号化圧縮する場合に比べて、音質の劣化を伴う。それに対し、必須符号と音質改善符号による階層化符号化を用いた場合は、再分析や再量子化の必要がないため、音質の劣化を最小限に抑えることが可能となる。

【0061】(4)他の実施形態  
上記各実施形態の説明においても、種々の変形実施形態について言及したが、以下に例示するような変形実施形態を挙げることができる。

【0062】上記各実施形態においては、パケット20

1は可変長としたが、固定長のパケットを用いることも可能である。ペイロード長が音声フレームのデータ長以下の場合は、複数のパケットにまたがり音声記録空間2に蓄積され、音声フレームのデータ長がペイロード長以下の場合は、1パケットに複数の音声フレームのデータが格納されて音声記録空間2に蓄積される。

【0063】音声記録空間については、パケットの種類や重要度等に応じて複数設けることも可能である。

【0064】音声要件の末尾を表すには、終端パケットを用いる他に、終端を示すフラグ等を符号化データが格納されている各パケットに付与しても良い。あるいは、各音声要件の開始パケットと終端パケットを管理する表を別途用意しても良い。

【0065】上記各実施形態では、直線的な記録空間を利用していたが、電子計算機で用いられる場合はファイルシステムの構成としても良い。これにより、記録空間の空領域の管理が容易となり、「ごみ集め」が不要となる。また、音声圧縮ブロックが生成した空領域を速やかに実時間記録ブロックが再利用できるようになり、空領域の有効利用が容易となる。

【0066】

【発明の効果】以上のように、本発明によれば、蓄積手段に蓄積された音声信号のうち再度符号化圧縮が必要なものについては圧縮の度合いと音声信号の重要度に応じて圧縮するようにしたので、蓄積手段の容量に応じて再生音質と蓄積容量の最適化を図ることができる。

【図面の簡単な説明】

【図1】第1の実施形態の全体構成を示すブロック図である。

【図2】従来の留守番電話装置の構成を示す図である。

【図3】第1の実施形態の音声記録区間の構成を示す図である。

【図4】第1の実施形態のパケットのヘッダの構成を示す図である。

【図5】第1の実施形態のパケットの型を示す表である。

【図6】第1の実施形態の非発声区間検出ブロックの構成を示す図である。

【図7】第1の実施形態の前方後方ハングオーバー回路の動作を示す図である。

【図8】第1の実施形態の音声圧縮ブロックの構成を示す図である。

【図9】第1の実施形態のフレーム重要度を設定するための構成を示す図である。

【図10】第1の実施形態の音声記録空間に書き込み及び指標読み出し指標を設けた場合の動作を示す図である。

【図11】第1の実施形態の圧縮参照表を示す図である。

【図12】第1の実施形態のフローチャートである。

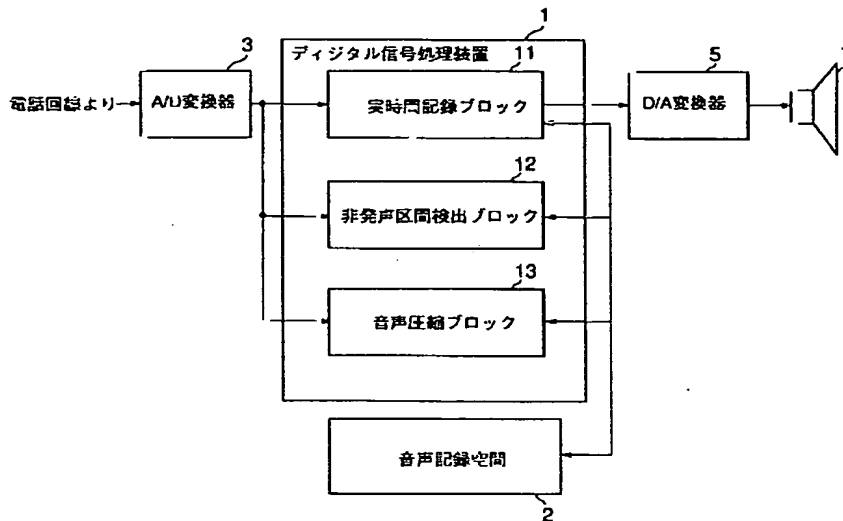
【図13】第2の実施形態の音声圧縮ブロックの構成を示す図である。

【図14】第3の実施形態の階層符号化の構成を示す図である。

【符号の説明】

1…デジタル信号処理装置、11…実時間記録ブロック、12…非発声区間検出ブロック、13…音声圧縮ブロック、2…音声記録空間、3…A/D変換器、5…D/A変換器、7…スピーカ。

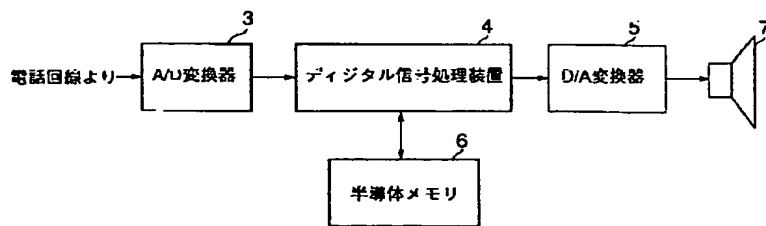
【図1】



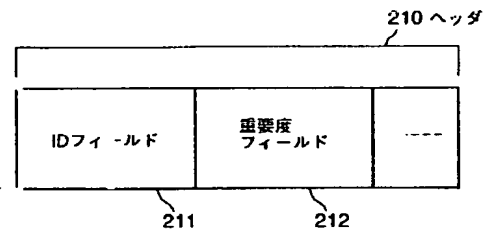
【図5】

ID	パケット内容
0	空
1	PCM符号
2	8kbps CELP
3	4kbps CELP
4	非発声
5	終端

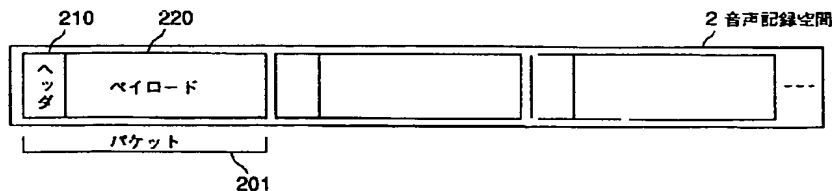
【図2】



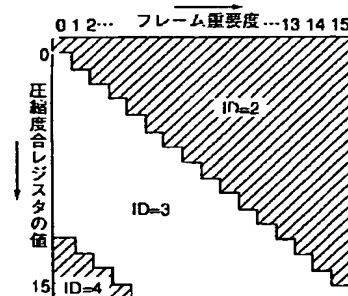
【図4】



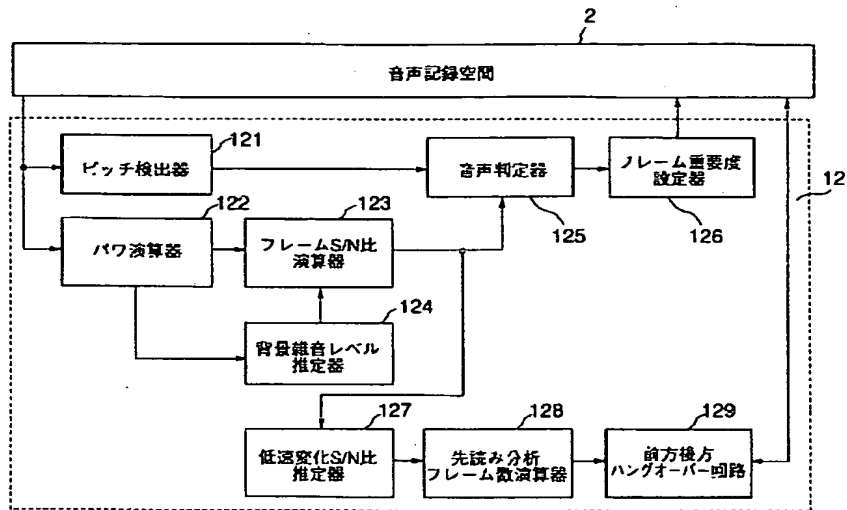
【図3】



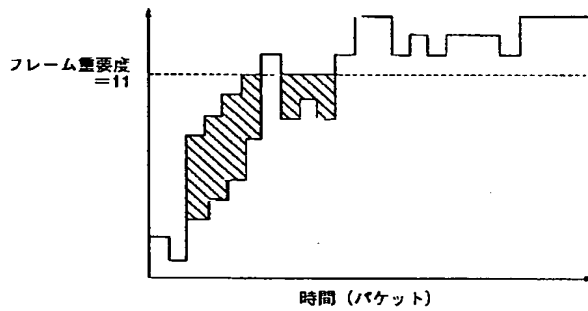
【図11】



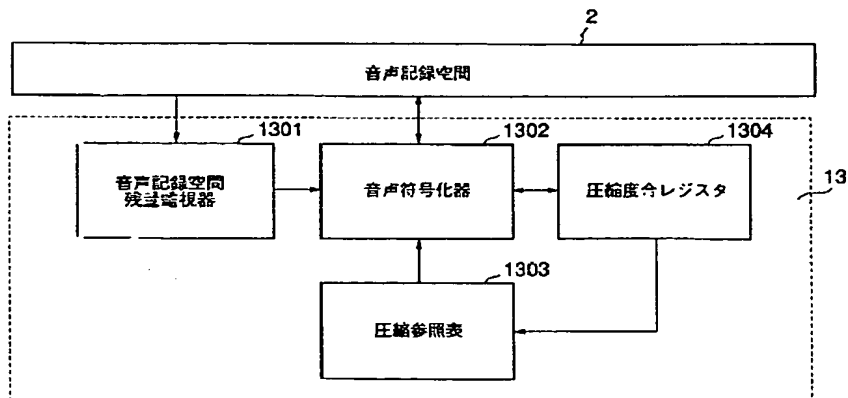
【図6】



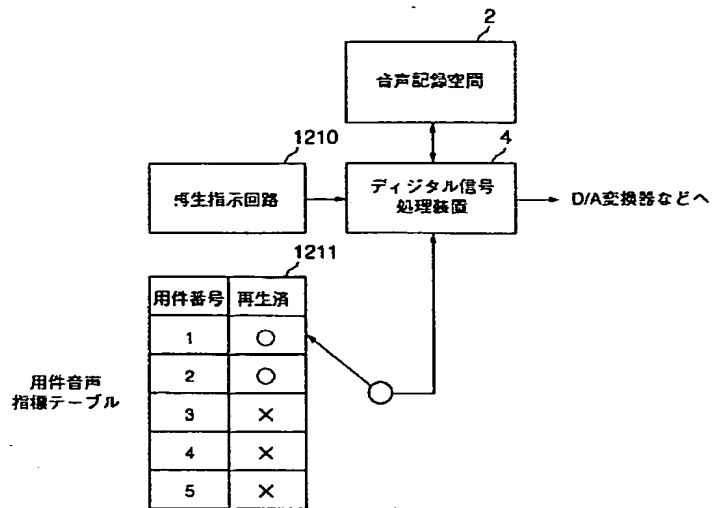
【図7】



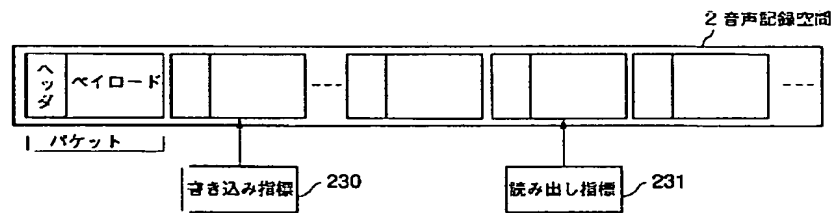
【図8】



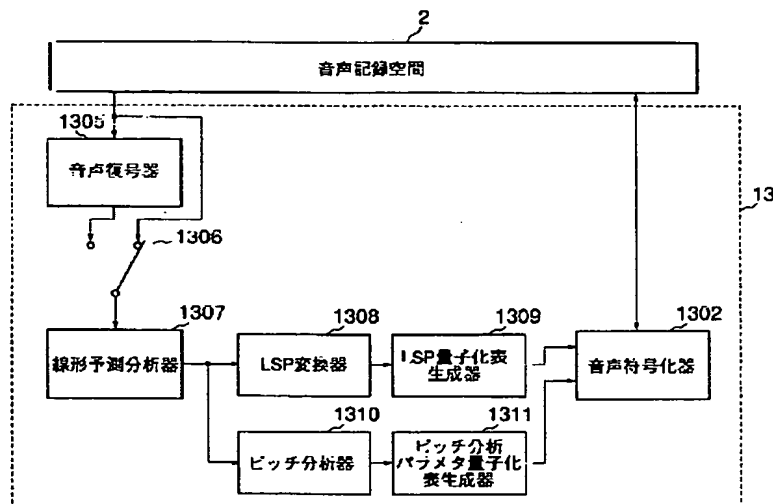
【図9】



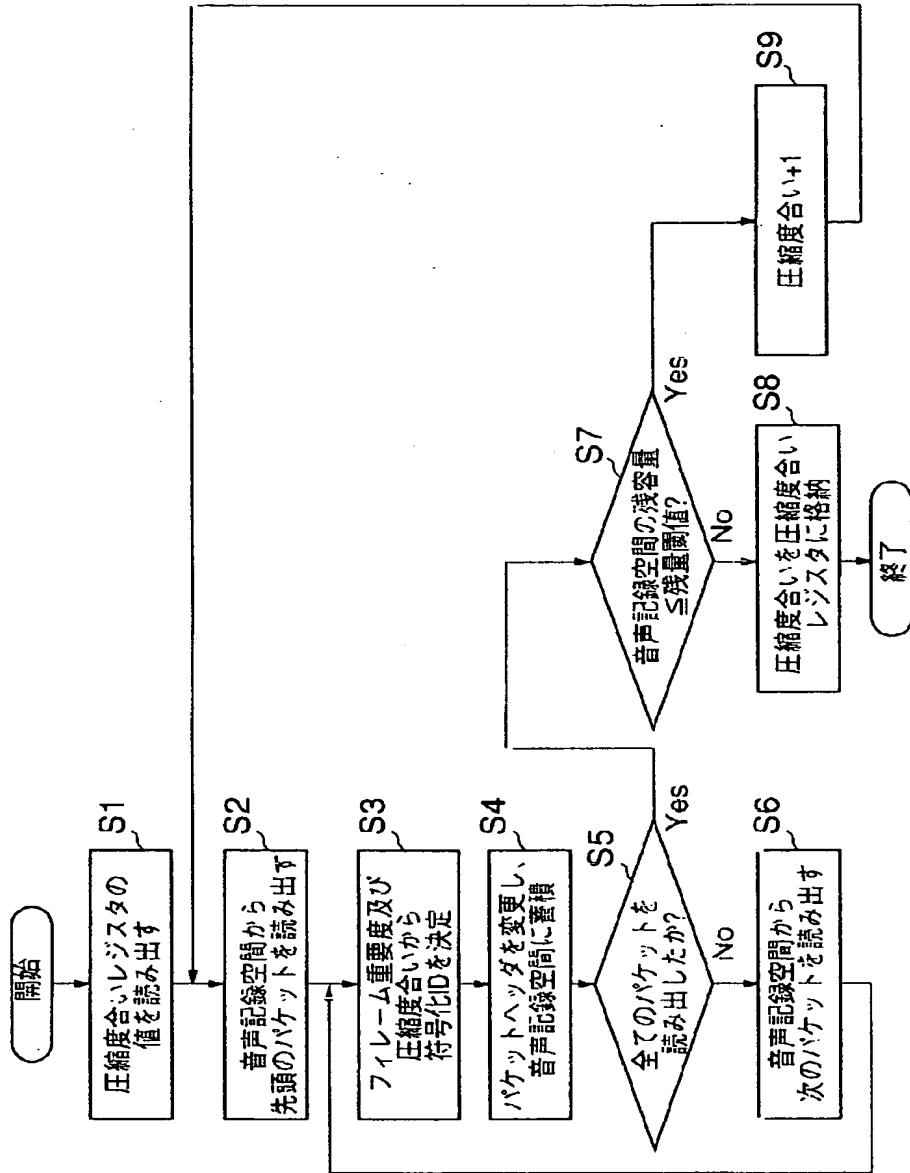
【図10】



【図13】



【図12】



【図14】

